# Exhibit 8

# Network Traffic Measurement for the Next Generation Internet

**Dimitrios P. Pezaros, B.Sc. (Hons.), MIEEE**

Computing Department
Lancaster University
England

SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

August 2005

Στη μνήμη των νεκρών της εξέγερσης του Πολυτεχνείου, τον Νοέμβρη 1973:

*Σπύρος Κοντομάρης* (57)
*Διομήδης Κομνηνός* (17)
*Σωκράτης Μιχαήλ* (57)
*Toril Margrethe Engeland* (22)
*Βασίλης Φάμελλος* (26)
*Γιώργος Σαμούρης* (22)
*Δημήτρης Κυριακόπουλος* (35)
*Σπύρος Μαρίνος* (31)
*Νίκος Μαρκούλης* (24)
*Αικατερίνη Αργυροπούλου* (76)
*Στέλιος Καραγεώργης* (19)
*Μάρκος Καραμανής* (23)

*Αλέξανδρος Σπαρτίδης* (16)
*Δημήτρης Παπαϊωάννου* (60)
*Γιώργος Γεριτσίδης* (47)
*Βασιλική Μπεκιάρη* (17)
*Δημήτρης Θεοδωράς* (5)
*Αλέξανδρος Βασίλης (Μπασρί) Καράκας* (43)
*Αλέξανδρος Παπαθανασίου* (59)
*Ανδρέας Κούμπος* (63)
*Μιχάλης Μυρογιάννης* (20)
*Κυριάκος Παντελεάκης* (44)
*Στάθης Κολινιάτης* (47)
*Γιάννης Μικρώνης* (22)

Στη μνήμη του *Νίκου Τεμπονέρα*

To the memory of the students and the civilians murdered during the tragic events that followed the public uprising at the National Technical University of Athens (NTUA) in November 1973:

*Spyros Kontomaris* (57)
*Diomidis Komninos* (17)
*Socrates Mihail* (57)
*Toril Margrethe Engeland* (22)
*Vasilis Famellos* (26)
*Giorgos Samouris* (22)
*Dimitris Kyriakopoulos* (35)
*Spyros Marinos* (31)
*Nikos Markoulis* (24)
*Ekaterini Argyropoulou* (76)
*Stelios Karageorgis* (19)
*Markos Karamanis* (23)
*Alexandros Spartidis* (16)
*Dimitris Papaioannou* (60)
*Giorgos Geritsidis* (47)
*Vasiliki Mpekiari* (17)
*Dimitris Theodoras* (5)
*Alexandros Vasilis (Bashri) Karakas* (43)
*Alexandros Papathanasiou* (59)
*Andreas Koumpos* (63)
*Michalis Myrogiannis* (20)
*Kyriakos Panteleakis* (44)
*Stathis Koliniatis* (47)
*Giannis Mikronis* (22)

To the memory of *Nikos Temponeras*

# Abstract

Measurement-based performance evaluation of network traffic is a fundamental prerequisite for the provisioning of managed and controlled services in short timescales, as well as for enabling the accountability of network resources. The steady introduction and deployment of the Internet Protocol Next Generation (IPNG-IPv6) promises a network address space that can accommodate any device capable of generating a digital heart-beat. Under such a ubiquitous communication environment, Internet traffic measurement becomes of particular importance, especially for the assured provisioning of differentiated levels of service quality to the different application flows. The non-identical response of flows to the different types of network-imposed performance degradation and the foreseeable expansion of networked devices raise the need for ubiquitous measurement mechanisms that can be equally applicable to different applications and transports.

This thesis introduces a new measurement technique that exploits native features of IPv6 to become an integral part of the Internet's operation, and to provide intrinsic support for performance measurements at the universally-present network layer. IPv6 Extension Headers have been used to carry both the triggers that invoke the measurement activity and the instantaneous measurement indicators in-line with the payload data itself, providing a high level of confidence that the behaviour of the real user traffic flows is observed. The in-line measurements mechanism has been critically compared and contrasted to existing measurement techniques, and its design and a software-based prototype implementation have been documented. The developed system has been used to provisionally evaluate numerous performance properties of a diverse set of application flows, over different-capacity IPv6 experimental configurations. Through experimentation and theoretical argumentation, it has been shown that IPv6-based, in-line measurements can form the basis for accurate and low-overhead performance assessment of network traffic flows in short time-scales, by being dynamically deployed where and when required in a multi-service Internet environment.

# Acknowledgments

There has been such a long list of people with whom interaction has enriched my life during the last five years that I could not possibly include all of them in these lines. I will always remember their invaluable contributions through the sharing of wonderful experiences, which influenced not only the evolution of this work, but also my overall personal blossoming. I am particularly indebted to the following people for their prompt and supportive encouragement, their guidance, and the selfless sharing of their knowledge and experience, without which this thesis would have not been possible as it stands.

I would like to express my very special gratitude to my academic supervisor, Professor David Hutchison, for his endless support, his professional guidance, his friendship, and the highly positive impact of his personality on my first steps in the international research community. By being my advisor throughout all my studies and my work at Lancaster University, David has taught me how to set high professional standards and conduct quality research in a self-confident and always optimistic manner.

Throughout my doctoral studies, I have been privileged to been offered an industrial fellowship from Agilent Laboratories, Scotland. I could not find the appropriate words to express my special gratitude to Agilent Technologies in general, and to individual members of the Telecommunications Solutions Department of Agilent Laboratories in particular, for their financial and practical support of this research. I wish this thesis stands up to their high standards. I would like to thank Professor Joe Sventek, who initiated this industrial fellowship, not only for his trust, but also for his highly-competent views and comments on numerous aspects of this work. This thesis would have not been possible as it stands, without the invaluable contributions of Dr. Francisco Garcia and Dr. Robert Gardner, with whom I had the privilege to work closely throughout my doctoral studies. I wish to sincerely thank them for always considering me as part of their extended team, for teaching me many aspects of systems and network research, and also for their friendship.

I would like to thank my colleagues in the Computing Department, Lancaster University, for their long-lasting support and encouragement. I am particularly grateful to Dr. Andrew Scott for being an excellent networking instructor, and also for practically supporting my work through the generous provision of equipment and infrastructural access that facilitated the real-world experimentation presented in this thesis. My very special thanks also go to Dr. Stefan Schmid for his invaluable support with the configuration of numerous experimental network topologies, and to Dr. Laurent Mathy for his patience and his continuous

encouragement during the last stages of my studies. I would like to thank Dr. Steven Simpson, Dr. Christopher Edwards, Dr. Paul Smith, Dr. Michael Mackay, and Dr. John Cushnie, for sharing their knowledge, but mostly for their friendship.

My very special thanks are also due to some very good friends who, over the years, offered their selfless emotional and also practical support. Panos Gotsis has generously shared his deep system administration knowledge, and provided invaluable systems support. Theodore Kypraios offered tremendous help during the descriptive statistical analysis of the experimental results documented in this thesis. Manolis Sifalakis has been a great friend and colleague whose axiomatic optimism taught me that modesty can perfectly match with self-confidence. Kostas Georgopoulos and Erasmia Kastanidi stood by me and offered great emotional support, especially during the last stages of this work, when suddenly everything seemed to be getting more difficult. I will never forget the Lancaster's good-old Greek Ph.D. gang and the great experiences I shared with Dr. Christos Efstratiou, Dr. Andrianos Tsekrekos, and Dr. Anthony Sapountzis.

My beloved Lena Gogorosi offered me an unquantifiable amount of love and support without which my life would have been very different. I feel that I could not thank her enough for the moments we shared together.

I wish to sincerely thank Deborah J. Noble for teaching me how to speak, read and write the English language during my childhood, and for remaining a good friend thereafter.

Finally but not least, my parents Pavlos Pezaros and Dorina Tsotsorou, and my grandparents Stelios and Soula Tsotsorou, have always been my main sources of encouragement and emotional strength. I could not thank them enough nor could I adequately express how much I owe them, for always being the wind beneath my wings throughout my entire life.

# Declaration

This thesis has been written by myself, and the work reported herein is my own. The documented research has been carried out at Lancaster University, and was fully funded by Agilent Technologies Laboratories, Scotland, through an industrial fellowship.

The work reported in this thesis has not been previously submitted for a degree in this, or any other form.

Dimitrios Pezaros, August 2005.

### 2.2.6 Bandwidth Estimation

Bandwidth estimation is a special case of active measurements, where synthetic traffic is injected into the network to try and characterise the amount of data that can be transferred by the infrastructure per unit of time. The area is lately seeing an increasing popularity and is sometimes considered to be exhibiting distinct characteristics from other measurement work[19]. This is not due to the use of specific infrastructures or certain protocols during the measurement process (as it was the case with most of the previous sections of this chapter), but mainly due to the focus being on the measurement practices and methodologies, and on assumptions (or lack of) that will produce accurate and unbiased results. Bandwidth is a fundamental property of a network connection, and producing a representative estimation of its metrics using raw packet values, requires an intensive investigation of measurement strategies that can minimise the heuristics and assumptions during the measurement process, as well as during the measurement analysis. This section briefly discusses the major issues and outlines the main measurement strategies in bandwidth estimation, which still remains a relatively new (sub-)area of network measurements research. Bandwidth measurements are viewed as complementary active probing techniques, yet the detailed analysis of the origins, theory, applications and implications of bandwidth estimation is beyond the scope of this thesis.

Within the data networks context, the term *bandwidth* quantifies the data rate that a network link or path can transfer. Three major metrics have been defined in the literature to identify different aspects of bandwidth. The *capacity* or *bottleneck bandwidth* of a link or path sets the upper limit on how quickly the network can deliver the sender's data to the receiver. The capacity $C$ of an $H$-hop end-to-end path is the maximum IP layer rate that the path can transfer from source to sink, and it depends on the underlying transmission technology and propagation medium [PrMD03]. The end-to-end capacity is determined by the minimum link capacity, i.e. the slowest forwarding element (*narrow link*) in the end-to-end chain that comprises the path.

$$C = \min_{i=0...H} C_i \tag{3}$$

Bottleneck bandwidth gives an upper bound on how fast a connection can *possibly* transmit data [Paxs97a]. The *available bandwidth* of a link relates to its 'spare' capacity during a certain time period, and relates not only on the underlying medium, but also on the traffic load. At any specific time instant, a link is either transmitting a packet at full link capacity or

[19] The First Bandwidth Estimation (BEst) workshop was organised by IETF's Internet Measurement Research Group (IMRG), CAIDA, and the US Department of Energy (DoE) in December 2003

it is idle, hence available bandwidth definition requires time averaging of the instantaneous utilisation over the time interval of interest [PrMD03]. The average utilisation $\bar{u}(t-\tau,t)$ for a period $(t-\tau,t)$ is given by

$$\bar{u}(t-\tau,t)=\frac{1}{\tau}\int_{t-\tau}^{t}u(x)dx \tag{4}$$

where $u(x)$ is the instantaneous utilisation of the link at time $x$. Hence, if $C_i$ is the capacity of a hop $i$ and $u_i$ is the average utilisation of that hop in the given time interval, then its average spare capacity is $C_i(1-u_i)$. The available bandwidth of an $H$-hop path is the minimum available bandwidth (*tight link*) of all $H$ hops [DoRM04].

$$A=\min_{i=0\ldots H}[C_i(1-u_i)] \tag{5}$$

Available bandwidth denotes how fast the connection *can* transmit while still preserving network stability, and never exceeds bottleneck bandwidth [Paxs97a]. Figure 2-8 shows the pipe model with fluid network traffic representation of a 3-link path, identifying the different notions of *capacity* and *available bandwidth* for each link. The figure also demonstrates that the *narrow link* ($C_1$) which determines the end-to-end capacity can be different from the *tight link* ($A_3$) which determines the end-to-end available bandwidth.

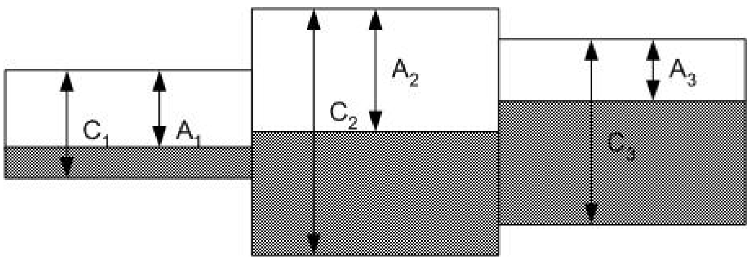


**Figure 2-8:** Pipe Model with Fluid Traffic of a Network Path

The third major bandwidth-related metric in TCP/IP networks is the throughput or *Bulk Transfer Capacity (BTC)* of a congestion-aware transport protocol (TCP) connection. However, as it has also been stated within the IPPM working Group [MaAl01], strictly defining the expected throughput of a TCP connection proves a challenging task, because it is influenced by numerous, non-static factors. These include the TCP transfer's size, the type of cross traffic (TCP or UDP), the number of competing TCP connections, the TCP socket buffer sizes at the sender and the receiver, the congestion along the reverse (ACK) path, the size of router buffers along the path, and the capacity and load of each link [PrMD03]. Within the bandwidth estimation community, BTC is used in coherence with the relevant IPPM metric specification (section 2.2.2) to denote *the maximum throughput obtainable by a single TCP connection* [MaAl01] whose ends implement all TCP congestion control algorithms [AlPS99].

BTC is TCP-specific and is fundamentally different from the available bandwidth metric, which is independent of any transport protocol, and it assumes that the average traffic load remains constant [PrMD03].

There currently are two major techniques for estimating *capacity* in individual hops and end-to-end paths, while newer deployments focus on the *available bandwidth* of Internet paths. These are mainly distinguished by *the way* they probe the network in order to estimate bandwidth, as opposed to *what type of traffic* they use.

- **Variable Packet Size (VPS) Probing**

VPS probing techniques try to measure the Round-Trip Time (RTT) from a source to each hop of a network path as a function of the probing packet size. They use the TTL field of the IP header to force probing packets to expire at a particular hop which will then generate the ICMP Time-Exceeded error message and send it back to the source. Upon reception of the ICMP message the source can measure the RTT, which consists of three delay components: the *serialisation delay* being the time (L/C) to transmit a packet of length L at a link of transmission rate C; the *propagation delay/latency* occurring due to the physical properties of the medium while transmitting each bit of a packet at a link and is independent of packet size; and the *queuing delay* occurring in the forwarding engine and the buffers of input and output ports of routers. By assuming a negligible serialisation delay for the small ICMP error packets, and also that, given a large number of probes one will eventually make the round trip with negligible queuing delays, VPS techniques compute the capacity of a hop as a linear function of the minimum RTT for a given probe packet size [Down99, PrMD03].

However, it has been lately suggested that VPS probing can cause consistent and significant underestimation of hop capacity, due to the presence of layer-2 store-and-forward devices (switches) that introduce additional latencies, not visible at (and hence non-computable by) layer-3 mechanisms [PrDM04, PrDM03].

- **Packet Pair/Train Dispersion**

Packet pair probing is used to measure the end-to-end capacity of a path. Multiple packet pairs consisting of two packets of the same size are sent back-to-back from a source to a receiver. The dispersion $\delta$ of a packet pair after a specific link of the path is the time distance between the complete transmission of the two packets [DoRM04].